



Estimación de la tasa de suicidio en Colombia por departamento para el período 2014-2019 mediante regresión beta clásica, semi-paramétrica y bayesiana.

Estimation of the suicide rate in Colombia by department for the period 2014-2019 through classical, semi-parametric, and Bayesian beta regression.

Michelle Vanesa Jiménez ^a
michellejimenez@usantotomas.edu.co

Dagoberto Bermudez Rubio ^b
dagobertobermudez@usta.edu.co

Resumen

Este trabajo de grado examina las tasas de suicidio en Colombia, desglosadas por departamentos, durante el periodo de 2014 a 2019. La estimación de estas tasas se realizó mediante modelos lineales generalizados y modelos aditivos generalizados, abordando aspectos de posición, escala y forma. Cabe destacar que la variable respuesta, la tasa de suicidio, supone una distribución beta, por lo que se emplearon específicamente GLM, GAMLSS y modelos bayesianos beta para su análisis. Además, se incorporaron modelos bayesianos para enriquecer el análisis con una perspectiva adaptativa basada en datos observados y previos. Se realizó la inclusión de información socioeconómica por departamento que permite una comprensión de estos factores en las variaciones de las tasas de suicidio a nivel regional. Se desea entender los determinantes contextuales de las tasas de suicidio en Colombia.

Palabras claves: Tasas de suicidio, modelos lineales generalizados, modelos bayesianos, distribución beta.

Abstract

This thesis examines suicide rates in Colombia, disaggregated by departments, during the period from 2014 to 2019. The estimation of these rates was conducted using generalized linear models and generalized additive models, addressing aspects of location, scale, and shape. It is noteworthy that the response variable, the suicide rate, assumes a beta distribution, so specifically, GLM, GAMLSS, and Bayesian beta models were employed for analysis. Additionally, Bayesian models were incorporated to enhance the analysis with an adaptive perspective based on observed and prior data. Socioeconomic information by department was included, allowing an understanding of these factors in the variations of suicide rates at the regional level. The aim is to comprehend the contextual determinants of suicide rates in Colombia.

Keywords: Suicide rates, generalized linear models, Bayesian models, beta distribution.

^a Estudiante Estadística Universidad Santo Tomás

^b Docente Estadística Universidad Santo Tomás

1 Introducción

El suicidio es considerado un problema de salud pública que ha incidido en las poblaciones desde el inicio de la civilización, siendo una práctica bastante común, se calcula que anualmente se presentan más de 700.000 suicidios al año (2021) [12], que en realidad podría ser una cifra mucho mayor considerando que existen países miembros de la Organización Mundial de Salud que no llevan a cabo este tipo de recolección de datos. Es considerado un problema de salud pública debido a: el impacto social en familia y amigos cercanos, los costos médicos que podrían ser destinados a otro tipo de enfermedades y principalmente es una conducta que se puede prevenir. La OMS ha destacado la necesidad de dar prioridad a esta causa de muerte, ya que existen herramientas oportunas que podrían contribuir a la mejora de estas estadísticas (2021) [12].

El fenómeno del suicidio presenta dos dimensiones: una centrada en factores personales y otra, factores sociales. En cuanto al ámbito personal, se incluyen elementos como antecedentes de comportamientos suicidas, diagnósticos psiquiátricos, género, grupos de edad, y otras variables. Estos aspectos individuales ofrecen un enfoque de la complejidad y diversidad de las experiencias que pueden contribuir al suicidio.

Por otro lado, los factores sociales y culturales desempeñan un papel fundamental en la comprensión del fenómeno del suicidio. La calidad de vida, la dinámica cultural y los factores económicos ejercen influencia sobre la vulnerabilidad de las personas. Por ejemplo, un entorno con condiciones económicas podría desfavorables podría o no crear obstáculos significativos para el crecimiento personal, social y académico de un individuo, contribuyendo así al riesgo de suicidio. En este sentido, se destaca la importancia de considerar no solo los aspectos individuales, sino también los contextos sociales más amplios al abordar la prevención y comprensión del suicidio.

La información fue recolectada a través de diferentes fuentes, entre ellas el Departamento Administrativo Nacional de Estadística (DANE) que recolecta información sobre las causas de defunción, proyecciones poblacionales, PIB real, etc, junto con 'Demographic and Health Surveys', encuestas que se realizan en numerosos países en conjunto con agencias gubernamentales y organizaciones internacionales de salud que recopilan datos demográficos, de salud y poblacionales sobre diferentes temas como salud, nutrición, ingresos, etc.

Los procesos de estimación se realizan mediante el modelo beta, que evalúa la media de la variable, la cual se define en función de los parámetros de media y dispersión (μ, ϕ), adicional mediante estimaciones GAMLSS (Generalized Additive Models for Location, Scale and Shape) que son modelos que se flexibilizan para permitir estimar, no solo la media, sino otros parámetros de la distribución de la variable respuesta, como funciones paramétricas lineales y/o aditivas no paramétricas de variables explicativas y/o efectos aleatorios. Para ajustar a los modelos se usa la estimación de máxima verosimilitud (penalizada). Por último, el enfoque bayesiano con un modelo beta proporciona un marco para modelar proporciones, permitiendo la incorporación de conocimientos previos y la actualización de creencias a medida que se observan los datos.

Este documento está organizado de la siguiente manera: La primera sección se dedica a desarrollar una introducción teórica acerca de la distribución beta, se exploran modelos semi-paramétricos que emplean splines penalizados, seguidos de modelos bayesianos. Además, se realiza una descripción de la base de datos y análisis descriptivo, luego, la aplicación de estos conceptos a través de múltiples modelos que tienen en cuenta variables socioeconómicas. Finalmente, se presentan los hallazgos más significativos derivados de este análisis junto con recomendaciones.

1.1 Planteamiento del problema

Existen diferentes investigaciones que han planteado el tema de suicidio, pero todo esto a través de factores sociales y personales, disminuye la cantidad de investigaciones cuando se desea evaluar la relación respecto a factores económicos.

Específicamente para Colombia los estudios que se han realizado toman en cuenta un solo departamento, tienen en cuenta otras variables como sexo, grupos de edad entre otras, solo toman en cuenta un factor

socioeconómico o bien son estudios realizados entre los años 1990 y 2010.

2 Justificación

La importancia de abordar el suicidio como un problema de salud pública radica en su impacto individual y social. La prevención del suicidio no solo implica intervenir en las causas individuales, sino también comprender y abordar los determinantes sociales y culturales que contribuyen a este fenómeno.

El suicidio en Colombia genera preocupación, y entender esta dinámica se vuelve crucial para desarrollar estrategias de prevención efectivas. La falta de datos exhaustivos y la necesidad de enfoques específicos que consideren la diversidad de factores involucrados hacen que esta investigación sea imperativa.

Además, la investigación propuesta no solo contribuirá a la comprensión de los factores que influyen en las tasas de suicidio en Colombia, sino que también los datos podrían ayudar en cuanto al diseño de intervenciones y políticas de prevención adaptadas a la realidad del país.

La investigación sobre el suicidio en Colombia es esencial para abordar un problema de salud pública crítico, proporcionando conocimientos valiosos para informar políticas y prácticas de prevención más efectivas y, en última instancia, reducir las tasas de suicidio y sus impactos asociados.

2.1 Objetivos

2.1.1 Objetivo general

Determinar si factores socioeconómicos influyen en las tasas de suicidio en Colombia por departamentos, para el periodo 2014 - 2019.

2.1.2 Objetivos específicos

- Identificar variables que inciden en las tasas de suicidio en Colombia.
- Analizar los resultados para cada uno de los modelos implementados.
- Seleccionar el mejor modelo que prediga la media de la tasa de suicidio.

2.2 Hipótesis

Existen factores socioeconómicos que inciden en las tasas de suicidio en Colombia.

3 Marco Teórico

En este capítulo, se explorarán las metodologías y definiciones esenciales necesarias para llevar a cabo las aplicaciones relacionadas con las tasas de suicidio.

3.1 Distribución beta clásica

El modelo de regresión beta fue introducida por Ferrari y Cribari Neto (2004) [6], es usualmente utilizada para modelar variables continuas (y), que asume el intervalo $(0,1)$ o cuando la variable toma el valor entre (a, b) donde $b > a$.

La variable aleatoria y se distribuye beta si su función de distribución está dada por:

$$f(y|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)y^{\alpha-1}(1-y)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}, \quad 0 < y < 1 \quad (1)$$

Donde $\alpha > 0$, $\beta > 0$ y $\Gamma(\cdot)$ es la función gamma. La media y varianza de y se define respectivamente como:

$$E[y] = \mu = \frac{\alpha}{\alpha + \beta} \quad (2)$$

$$Var[y] = \sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \quad (3)$$

Algunos ejemplos donde la variable asume esta distribución puede ser la proporción de personas que votan por 'x' candidato en elecciones, tasa de desempleo, rendimiento de un estudiante en alguna materia.

De manera gráfica y utilizando diferentes parámetros para alpha (α) y beta (β) podemos evidenciar la función de densidad de probabilidad así:

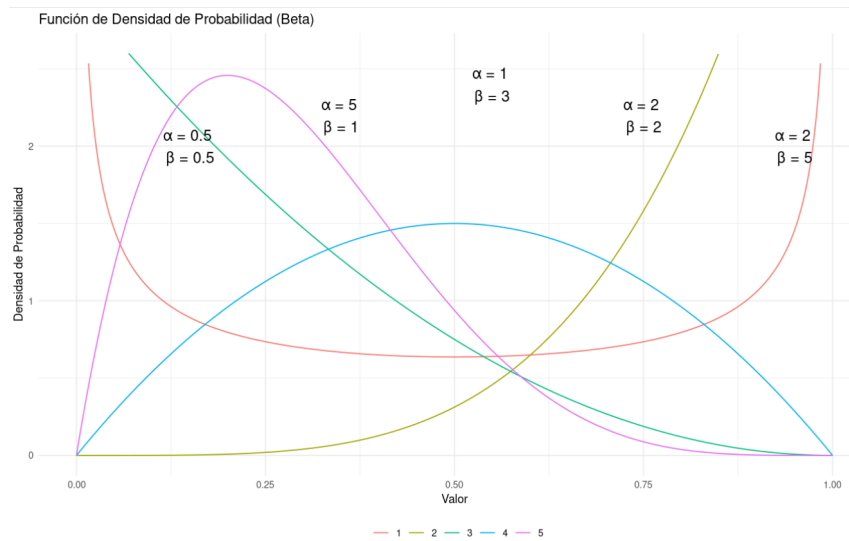


Figure 1: Función de densidad de probabilidad beta

Sin embargo, para realizar el análisis respectivo es útil modelar la media de la variable respuesta, Ferrari y Cribari Neto (2004) [6] propusieron una parametrización diferente utilizando:

$$\mu = \frac{\alpha}{\alpha + \beta}, \quad \phi = \alpha + \beta, \quad \alpha = \mu\phi, \quad \beta = (1 - \mu)\phi \quad (4)$$

3.1.1 Distribución beta para la media y varianza

La distribución beta es muy útil para poder evaluar proporciones, tasas sin embargo para este caso se evaluará la media de la variable, se utiliza un parámetro de precisión. La función de densidad se define ahora como:

$$f(y|\mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1 \quad (5)$$

Donde $0 < \mu < 1$ y $\phi > 0$, denotando $y \sim \mathcal{B}(\mu, \phi)$, $E(y) = \mu$ y $Var(y) = \sigma^2 = \frac{\mu(1-\mu)}{1+\phi}$. El parámetro ϕ se conoce como el parámetro de precisión, con μ conocido, cuando mayor sea ϕ menor será la varianza de y ; ϕ^{-1} se define como un parámetro de dispersión.

Sea $y_1, y_2, \dots, y_n$ una muestra aleatoria tal que $y_i \sim \mathcal{B}(\mu, \phi)$, $i = 1, \dots, n$. El modelo de regresión beta se define como:

$$g(\mu_i) = x_i^\top \beta, \quad (6)$$

$$h(\phi_i) = z_i^\top \gamma \quad (7)$$

Donde $\beta = (\beta_1, \dots, \beta_k)$ y $\gamma = (\gamma_0, \dots, \gamma_k)$ son respectivamente los vectores de los parámetros: media y precisión, x_i y z_i son respectivamente conjuntos de variables explicativas para la media y la precisión para la i -ésima observación.

En Cepeda-Cuervo (2001) [3] se propuso que g es una función logit y h una función exponencial, pero otros enlaces pueden ser asumidos.

3.2 Modelo de regresión beta semiparamétrico

Los modelos aditivos generalizados para posición, escala y forma (GAMLSS), son modelos de regresión semi-paramétricos. Paramétricos en cuanto a que requieren asumir que la variable respuesta siga una determinada distribución de probabilidad paramétrica (normal, beta, gamma, etc) y semi porque los parámetros pueden ser modelados, cada uno de ellos como funciones no paramétricas (lineales, aditivas o no lineales) (2008) [11]. Con esta versatilidad los modelos GAMLSS es una herramienta útil cuando las variables siguen varias distribuciones (no normales, asimétricas, con varianza no constante, etc).

Los modelos GAMLSS asumen que la variable dependiente tiene una función de densidad definida por cuatro parámetros (μ, σ, ν, τ) que respectivamente determinan su posición (media), escala (desviación estándar), y forma (asimetría y curtosis). Cada uno de estos puede variar independientemente de los otros en función de los predictores.

Este tipo de modelos estiman cuatro funciones, donde cada uno establece la relación entre las variables predictoras y uno de los parámetros, existen muchas distribuciones que se puede emplear para la variable respuesta (y), $f(y_i|\mu_i, \sigma_i, \nu_i, \tau_i)$, con la condición que el logaritmo de su función de densidad y su primera derivada respecto a cada uno de los parámetros puedan calcularse. El modelo GAMLSS se escribe de la siguiente manera:

$$y \sim D(\mu, \sigma, \nu, \tau), \quad (8)$$

$$\eta_1 = g_1(\mu) = x_1\beta_1 + s_{11}(x_{11}) + \dots + s_{1J_1}(x_{1J_1}), \quad (9)$$

$$\eta_2 = g_2(\sigma) = x_2\beta_2 + s_{21}(x_{21}) + \dots + s_{2J_2}(x_{2J_2}), \quad (10)$$

$$\eta_3 = g_3(\nu) = x_3\beta_3 + s_{31}(x_{31}) + \dots + s_{3J_3}(x_{3J_3}), \quad (11)$$

$$\eta_4 = g_4(\tau) = x_4\beta_4 + s_{41}(x_{41}) + \dots + s_{4J_4}(x_{4J_4}) \quad (12)$$

Donde $D(\mu, \sigma, \nu, \tau)$ es una distribución de cuatro parámetros, donde ν y τ son parámetros de forma que a menudo son relacionados con la asimetría y curtosis de la distribución. Este modelo fue introducido por Rigby y Stasinopoulos (2005) [10].

- Distribuciones: La forma de la distribución $D(\mu, \sigma, \nu, \tau)$ es genérica e implica que debe ser una forma paramétrica.
- Términos aditivos: Las variables explicativas pueden afectar los parámetros de la distribución especificada en diferentes formas. Se permite utilizar funciones paramétricas lineales o no lineales, funciones de suavizado no paramétricas de variables explicativas, existen varios términos aditivos como: p-splines (B-splines penalizados), p-splines monótonos, p-splines de ciclo, p-splines de coeficiente variable, entre otros.

Los splines penalizados fueron introducidos por Eilers y Marx (1996) [5], son polinomios por tramos definidos por funciones de base B-spline en la variable explicativa, donde los coeficientes de las funciones de base están penalizados para garantizar suficiente suavidad (ver Eilers y Marx 1996) [5].

La idea principal es combinar la flexibilidad de los B-splines con la penalización para evitar el sobreajuste y mejorar la estabilidad numérica. Dentro de la penalización se agrega un término de penalización a la función de ajuste para controlar su suavidad. La penalización se aplica a las derivadas de alto orden de la función, y su intensidad está controlada por un parámetro de penalización.

El objetivo es encontrar la función de B-spline que minimiza una función de pérdida, que incluye la suma de los cuadrados de las diferencias entre los valores observados y los predichos, así como los términos de penalización.

Características:

- Amplia gama de distribuciones para la variable respuesta.
- Todos los parámetros de la distribución pueden modelarse como funciones de variables explicativas.
- Variedad de términos aditivos en los modelos para los parámetros de distribución.
- Amplía los modelos estadísticos básicos permitiendo el modelado flexible con sobre - dispersión, exceso de cero, asimetría y curtosis en los datos.
- El algoritmo ajustado es modelar, donde diferentes componentes se pueden agregar fácilmente.

3.3 Modelo de regresión beta bayesiana

Cepeda - Cuervo (2001)[3] y Cepeda et.al (2005)[1] presentan el método bayesiano y el método de Monte Carlo basado en cadenas de Markov (MCMC).

Para poder aplicar un enfoque bayesiano para ajustar un modelo de regresión beta, es necesario ajustar los parámetros del modelo, para ello es necesario especificar las distribuciones previas para los parámetros.

Por simplicidad se asume que las a priori son funciones basados en conocimientos previos: $\beta \sim N(\mathbf{b}, \mathbf{B})$ y $\gamma \sim N(\mathbf{g}, \mathbf{G})$. La función a posteriori se define como:

$$P(\beta, \gamma | data) \propto P(data | \beta, \gamma) \cdot P(\beta, \gamma) \quad (13)$$

La distribución a posteriori es analíticamente intratable debido a que no se puede expresar de manera analítica en una forma cerrada o con una fórmula matemática simple, además no es sencillo generar el muestreo aleatorio, Cepeda (2001) [3] propone muestras aleatorias utilizando un proceso iterativo para β y γ de las distribuciones condicionales a posteriori, $P(\beta | \gamma, data)$ y $P(\gamma | \beta, data)$, respectivamente. Pero dado que estas distribuciones son analíticamente intratables, la metodología bayesiana propuesta por Cepeda (2001)[3] Cepeda and Gammerman (2005)[1] donde es necesario construir núcleos de transición normales (q_1, q_2), este es aplicable.

Para construir las funciones es necesario definir variables de observación de trabajo para aproximarnos $h(\mu_i)$ y $g(\phi_i)$ alrededor de los valores actuales de μ y ϕ respectivamente. Se encuentra definido como el primer orden de Taylor aproximado a las funciones reales $h(t_1)$ y $g(t_2)$, donde t_1 y t_2 son variables aleatorias apropiadas tal que $E(t_1) = \mu$ y $E(t_2) = \phi$. De este modo $E(t_1) = \mu$ para $t_1 = Y$, si el promedio del modelo esta dado por $(g(\mu_1) = x_i^\top \beta)$, la variable observacional está dada por:

$$\tilde{y} = x_i^\top \beta^{(c)} + \frac{y_i - \mu_i^{(c)}}{(\mu_i^{(c)})(1 - \mu_i^{(c)})}, \quad i = 1, 2, \dots, n, \quad (14)$$

$\mu_i^{(c)}$ y $\beta^{(c)}$ son los valores de μ y β , asumiendo que $\tilde{y}_i, i = 1, \dots, n$ tiene una distribución normal y asumiendo que la distribución a priori condicional se distribuye así: $\beta | \gamma \sim N(b, B)$ la función de kernel q_1 está dada por la distribución a posteriori obtenida de la combinación de la distribución a priori con el modelo de observación trabajado $\tilde{y} \sim N(x_i^\top \beta, \tilde{\sigma}_i^2)$, donde $\tilde{\sigma}_i^2 = Var(\tilde{y}_i)$, es decir por:

$$q_1(\beta | \beta^{(c)}, \gamma^{(c)}) = N(\mathbf{b}^*, \mathbf{B}^*), \quad (15)$$

$$\mathbf{b}^* = \mathbf{B}^*(\mathbf{B}^{-1}\mathbf{b} + \mathbf{X}^\top \Sigma^{-1} \tilde{Y}) \quad \mathbf{B}^* = (\mathbf{B}^{-1} + \mathbf{X}^\top \Sigma^{-1} \mathbf{X})^{-1} \quad (16)$$

Donde Σ es una matriz diagonal, con entradas diagonales $\tilde{\sigma}_i^2, i = 1, 2, \dots, n$ de la variable de trabajo Y .

Para ϕ se define la variable de trabajo $\tilde{y}_i$ en términos de $h(\cdot)$ de esta manera se construye una variable aleatoria auxiliar cumpliendo $E(\tilde{y}_i) = \phi$. La variable se define en términos de Y_i , los parámetros μ y ϕ . Luego, se define la estimación de parámetros para ϕ , en función de Z_i

$$\tilde{y}_i = z_i^\top \gamma^{(c)} + \frac{y_i}{\mu_i} - 1, \quad i = 1, 2, \dots, n \quad (17)$$

De igual forma se tiene que $\tilde{y}_i$ se distribuye de manera normal y la función de distribución a priori condicional para γ es $\gamma | \beta \sim N(\mathbf{g}, \mathbf{G})$, la función de kernel q_2 está dada por la distribución posteriori obtenida de la combinación de la distribución priori con el modelo de observación trabajado $\tilde{y}_1 \sim N(z_i^\top \gamma, \tilde{\sigma}_i^2)$, donde $\tilde{\sigma}_i^2 = Var(\tilde{y}_i)$, es decir por:

$$q_2(\gamma|\gamma^{(c)}, \beta^{(c)}) = N(\mathbf{g}^*, \mathbf{G}^*), \quad (18)$$

$$\mathbf{g}^* = \mathbf{G}^*(\mathbf{G}^{-1}\mathbf{g} + \mathbf{Z}^\top \psi^{-1} \tilde{Y}) \quad G^* = (\mathbf{G}^{-1} + \mathbf{Z}^\top \psi^{-1} \mathbf{Z})^{-1} \quad (19)$$

,

ψ es una matriz diagonal, con entradas diagonales $\tilde{\sigma}_i^2$ para $i = 1, 2, \dots, n$ de la variable de trabajo Y .

Con los kernels obtenidos en (11) y (14), los componentes β y γ de (β_0, γ_0) , se seguirán los siguientes pasos:

1. Inicia el contador $j = 1$ y establecer valores iniciales (β_0, γ_0) a $(\beta, \gamma)^\top$
2. Mover el vector β a un nuevo valor de ψ generado por la densidad $q_1(\beta^{(j-1)}, \cdot)$.
3. Calcular la probabilidad de aceptación de movimiento, $\alpha(\beta^{(j-1)}, \psi)$, si es un valor aceptado, entonces $\beta^{(j)} = \psi$, si no es aceptado entonces $\beta^{(j)} = \beta^{(j-1)}$.
4. Mover el vector γ a un nuevo valor de ψ , generado y propuesto por la densidad $q_2(\gamma^{j-1}, \cdot)$
5. Calcular la probabilidad de aceptación de movimiento, $\alpha(\gamma^{(j-1)}, \psi)$, si es un valor aceptado entonces $\gamma^{(j)} = \psi$, si no es aceptado entonces $\gamma^{(j)} = \gamma^{(j-1)}$.
6. Por último, cambiar el contador de j a $j + 1$ e ir a el paso dos hasta obtener convergencia.

3.4 Criterios de información

Debido a que se generan diferentes modelos a través de criterios que permitan realizar la comparación y calidad del ajuste del modelo. El AIC (Akaike Information Criterion), definido en Spiegelhalter et al. (2002) [4], junto con el criterio DIC (Deviance Information Criterion) basado en la distribución a posteriori de la estadística deviance $D(\theta)$:

$$D(\theta) = -2\log(p(Y|\theta)) + 2\log(f(Y)), \quad (20)$$

$f(Y)$ denota la distribución marginal de Y .

El primer término de DIC mide la bondad de ajuste, el segundo término penaliza la complejidad del modelo, calculado como:

$$DIC = E_{\theta|Y}(D(\theta + p_D)) \quad (21)$$

En la ecuación, el primer término representa la esperanza a posteriori del deviance, mientras que el segundo término refleja el número efectivo de parámetros en el modelo. Este criterio resulta especialmente valioso en situaciones de selección de modelos donde las distribuciones a posteriori se obtienen mediante simulaciones mediante MCMC. Los modelos que ofrecen los mejores ajustes son aquellos que presentan valores menores en el DIC.

4 Aplicación

4.1 Análisis descriptivo de los datos

La base fue construida teniendo en cuenta los datos obtenidos y recolectados por el DANE y contiene información de veinticuatro departamentos, debido a la disponibilidad de la información de los treinta y tres departamentos (incluido Bogotá) no fueron tenidos en cuenta: Amazonas, Arauca, Archipiélago de San Andrés y Providencia y Santa Catalina, Casanare, Guainía, Guaviare, Putumayo, Vaupés y Vichada desde el 2014 hasta el 2019.

La información de las siguientes variables se encuentran anualizadas por departamentos, se tuvieron en cuenta:

- Defunciones anuales por la causal: 'Lesiones autoinfligidas intencionalmente (suicidios) y secuelas'.
- Proyecciones de la población en Colombia por departamento y año.
- **PIB a total precios constantes (año de referencia 2015):**

Es el valor total de los bienes y servicios finales producidos por el país durante un periodo definido. Se utiliza el 'a precios constantes' pues es una medida económica que ajusta los valores del PIB para tener en cuenta los cambios en los precios con el tiempo.

- **Índice de pobreza monetaria:**

Foster, Greer y Thorbecke (FGT) (1984)[7] definieron ciertos parámetros donde se definen los indicadores de pobreza monetaria: la incidencia, la brecha y la severidad. La incidencia de la pobreza monetaria mide el porcentaje de la población que tiene un ingreso per cápita de la unidad de gasto por debajo de la línea de pobreza según el dominio geográfico (Nacional o departamental).

La línea de pobreza es el costo per cápita mínimo de una canasta básica de bienes y servicios (alimentarios y no alimentarios) en un área geográfica determinada.

Para poder medir estos parámetros (incidencia, brecha y severidad) se deben establecer umbrales que determinen personas en condición de pobreza. El Departamento Administrativo Nacional de Estadística toma la formula por parte de Foster, Greer y Thorbecke (FGT) (1984) [7]:

$$P_{\alpha}(y; z) = \frac{1}{n} \sum_{i=1}^q \left(\frac{z - y_i}{z} \right)^{\alpha} \quad (22)$$

Donde z hace referencia a la línea de pobreza, y_i toma en cuenta el ingreso per cápita del hogar del individuo, n la población, q es el número de personas total cuyos sus ingresos se encuentran por debajo de la línea de pobreza z , $\alpha = 0$ indica la pobreza monetaria, $\alpha = 1$ mide la brecha de pobreza y $\alpha = 2$ mide la severidad de la pobreza.

- **Coefficiente de Gini**

El coeficiente de Gini es un indicador de desigualdad que mide la concentración de la riqueza en un área geográfica determinada. El coeficiente de Gini mide la distancia entre la curva de Lorenz y la línea de equidistribución, dependiendo la distancia entre estas, el coeficiente de Gini tomara valores entre cero y uno, donde cero representan igualdad absoluta y uno desigualdad absoluta.

- **Índice de Desarrollo Humano**

Es una medida resumida que promedia en dimensiones clave del desarrollo humano: salud, educación y nivel de vida (esperanza de vida - tasa de alfabetización, PIB), poseer conocimientos y disfrutar

de un nivel de vida decente. El IDH es la media geométrica de los índices normalizados para cada una de las tres dimensiones.

- **Índice de salud**

El índice de salud utiliza como indicador una de las secciones del IDH, es decir esperanza de vida al nacer, es decir la media de años de cierta población.

- **Índice de ingresos**

El índice de ingresos es el logaritmo natural del Ingreso nacional bruto per-capital en dólares, a nivel sub nacional se estimaron en función de la riqueza de los hogares. Para ello se utilizó el Índice de Riqueza Internacional (IWI), que mide la riqueza de los hogares a partir de información sobre propiedad de activos, calidad de la vivienda y acceso a servicios públicos.

Se identificó el tipo de departamentos con las mayores tasas de suicidio, tomando como referencia el año 2019, representando las variaciones del ranking. Vaupés es el departamento que, exceptuando el 2019 ha ocupado el puesto número uno. A pesar de que Risaralda, en años anteriores no ocupa los primeros puestos, se presenta un aumento en la tasa en el año más reciente. Por otro lado, Arauca ha sido el departamento que ha bajado de puestos en el Ranking.

Top	Departamento	2019	2018	2017	2016	2015	2014
1	Amazonas	1	3	3	10	2	4
2	Vaupés	2	1	1	1	1	1
3	Risaralda	3	7	11	12	7	9
4	Tolima	4	8	5	7	6	5
5	Arauca	5	4	2	2	3	3
6	Caldas	6	5	15	9	5	10

Table 1: Top tasas de suicidio por departamentos, 2014-2019.

Por año podemos ver de manera gráfica el comportamiento de la tasa de suicidios, presentando un comportamiento similar año tras año:

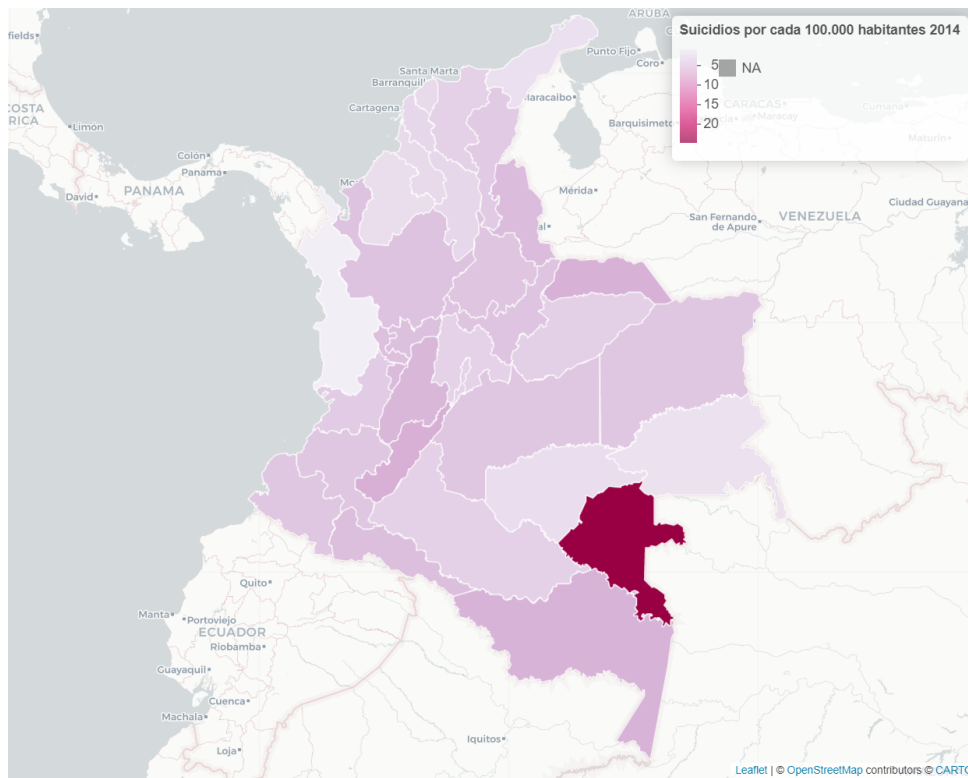


Figure 2: Tasa de suicidios 2014

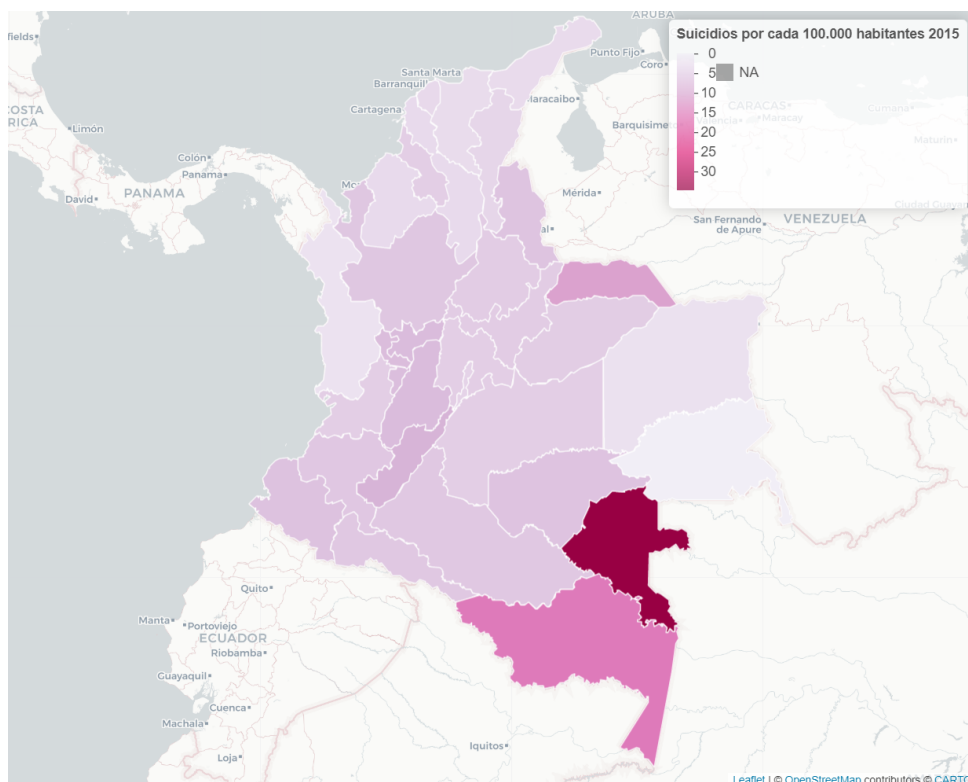


Figure 3: Tasa de suicidios 2015

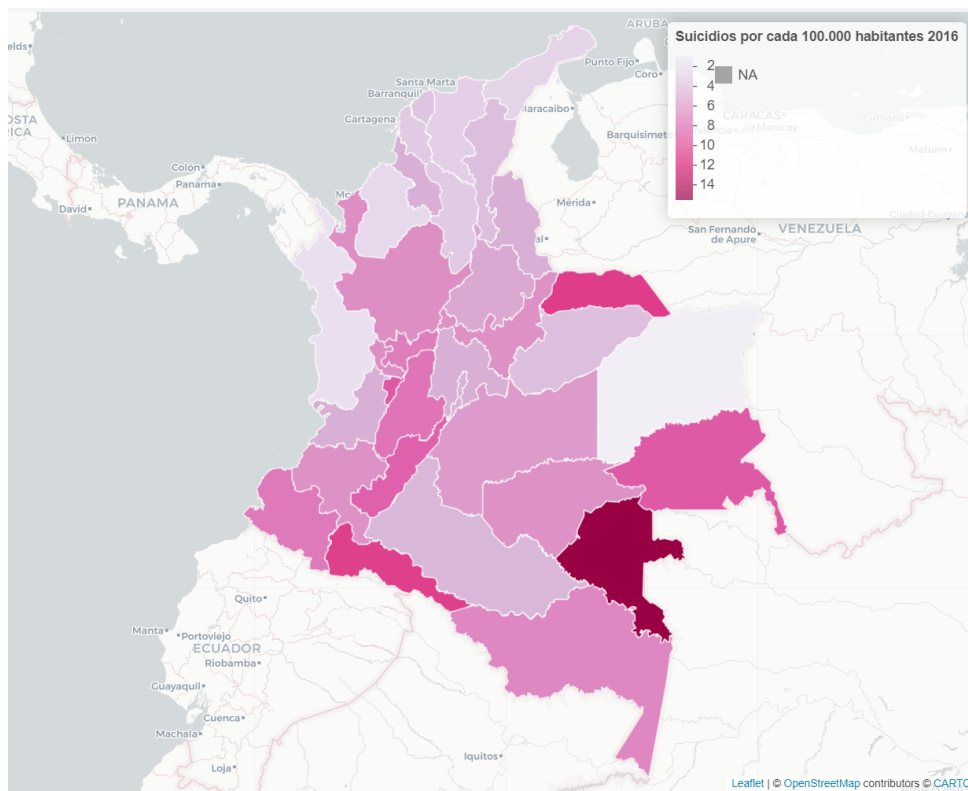


Figure 4: Tasa de suicidios 2016

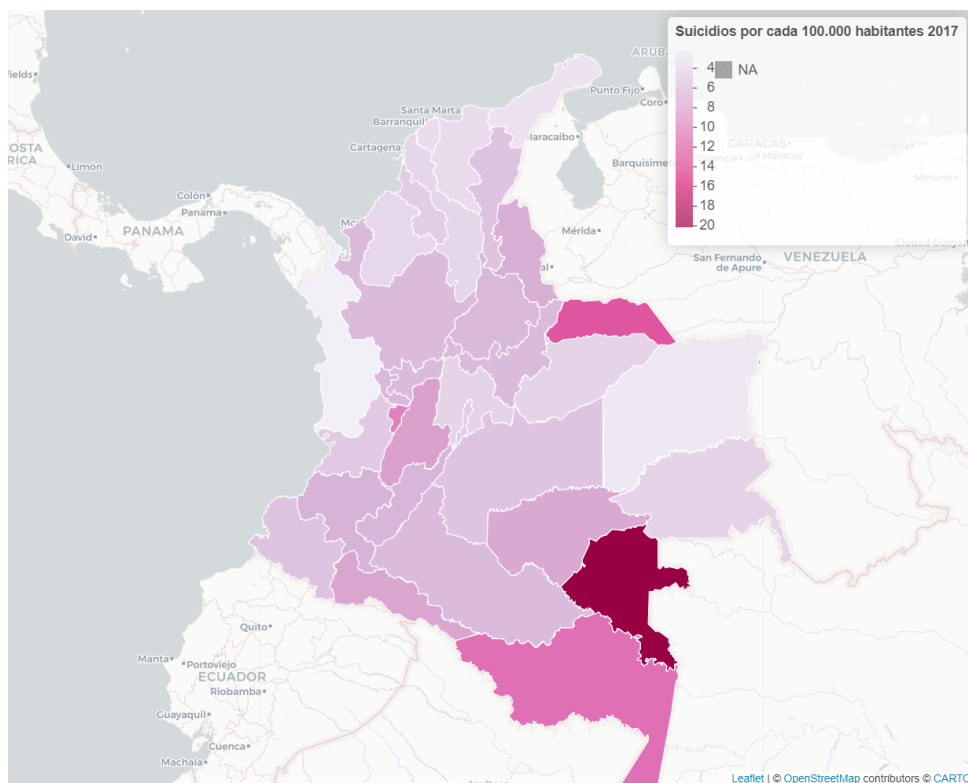


Figure 5: Tasa de suicidios 2017

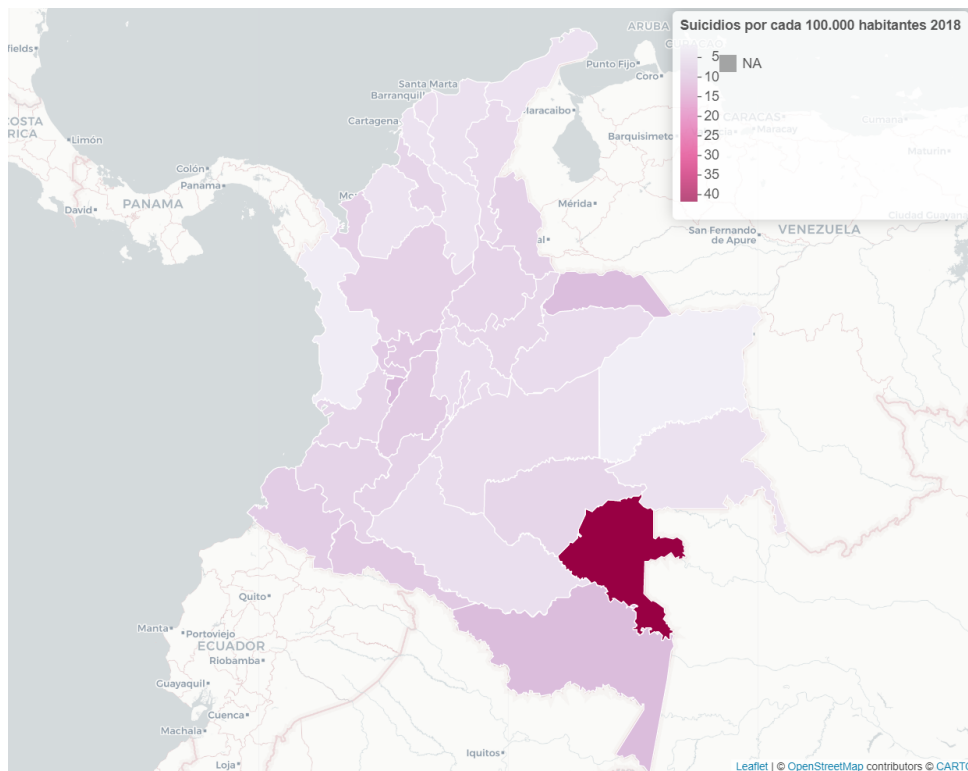


Figure 6: Tasa de suicidios 2018

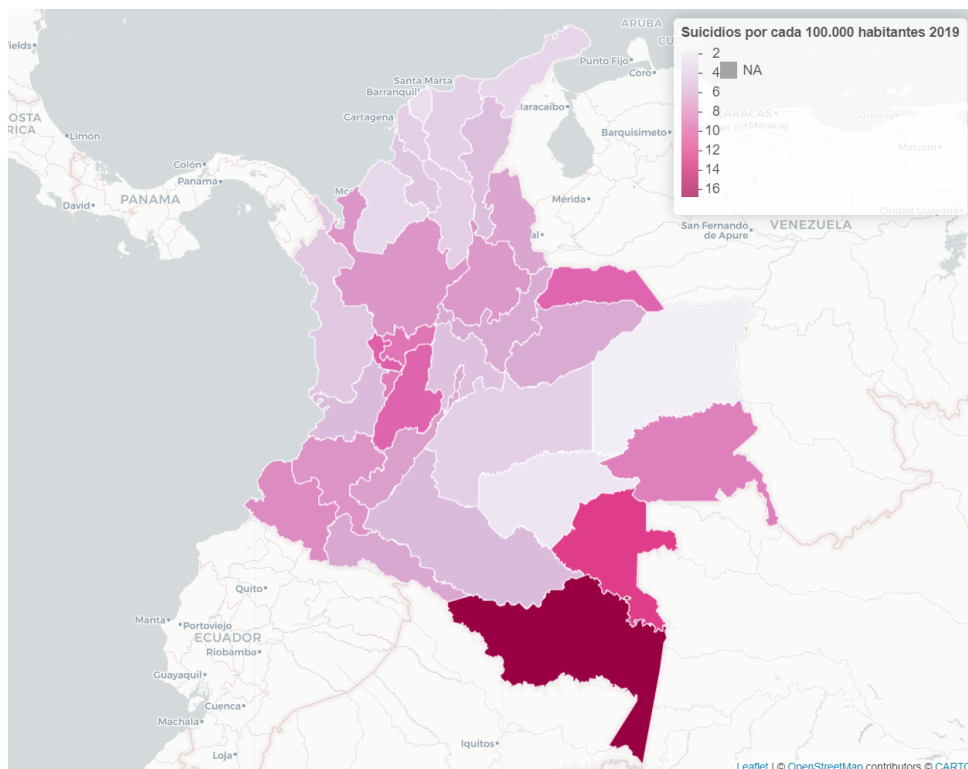


Figure 7: Tasa de suicidios 2019

La zona cafetera presentó una mayor incidencia de suicidios, respecto al resto. Los casos de suicidio en Vaupés continuán siendo una cifra alarmante, solo para el año 2018 hubo 42 suicidios por cada 100.000 habitantes.

A través de los años el rango de las tasas de suicidio varia y no es creciente pues existen años donde las tasas más bajas se ubican en 2019 y 2016, así como tasas altas para el año 2015 y 2018, como se aprecia en la tabla número uno, la tasa a través del tiempo en los departamentos crece y decrece.

4.2 Construcción de modelos

4.2.1 Correlación de variables

En la estimación de modelos y debido a la disponibilidad de los datos únicamente se trabajó con 24 departamentos: Antioquia, Atlántico Bogotá, Bolívar, Boyacá, Caldas, Caquetá, Cauca, Cesar, Córdoba, Cundinamarca, Chocó, Huila, Guajira, Magdalena, Meta, Nariño, Norte de Santander, Quindío, Risaralda, Santander, Sucre, Tolima y Valle del Cauca.

Para estructurar el modelo se evaluaron las correlaciones entre las variables elegidas para la construcción del modelo. Por ende, se evaluó la multicolinealidad de los datos.



Figure 8: Correlación entre variables



Figure 9: Correlación entre variables

Se evidencia una correlación alta en los índices de pobreza monetaria e índice de ingresos, por lo tanto, no fueron tenidas en cuenta para la construcción de los modelos.

4.2.2 Modelo de regresión beta clásico

Utilizando la función GAMLSS de Rstudio, se ajustan los modelos de regresión beta que por defecto utiliza la función de enlace "logit" para la media.

Se realizó la comparación de los modelos a través de los criterios de información Akaike Information Criterion (AIC) junto con 'Bayesian Information Criterion' combinando diferentes variables y/o suavización por splines penalizados.

Modelo	Variables del modelo	AIC	SBC
Mod. uno	Intercepto, Departamento, Año, PIB, IDH, Health, Desempleo, Gini	-987.03	-883.08
Mod. dos	Intercepto, Departamento, PIB, IDH, Health, Desempleo, Gini	-994.84	-905.74
Mod. tres	Intercepto, Departamento, PIB, IDH, Health, Desempleo	-996.63	-910.50
Mod. cuatro	Intercepto, Departamento, PIB, IDH, Desempleo	-997.23	-914.08

Table 2: Modelos beta clásicos ajustados a la tasas de suicidio en Colombia.

Comparando los criterios de información el mejor modelo que se ajuste es el número cuatro:

$$\text{logit}(\mu_i) = \hat{\beta}_0 + \hat{\beta}_1 \text{Departamentos}_i + \hat{\beta}_2 \text{PIB}_i + \hat{\beta}_3 \text{IDH}_i + \hat{\beta}_4 \text{Desempleo}_i \quad (23)$$

Sobre este modelo se realiza la interpretación del modelo número cuatro, mostrando las variables significativas (el departamento base fue Antioquia).

Parámetro	Estimador	Std.Error	z.Value	Pr (> t)	
Intercepto	-14.21	1.503	-9.454	4.22e-16	***
DepAtlantico	-0.2631.	0.4229	-0.622	0.535085	
DepBogota	-1.583	0.4205	-3.764	0.000263	***
DepBolívar	0.2229	0.4224	0.528	0.598728	
DepBoyaca	0.7568	0.4597	1.647	0.102332	
DepCaldas	0.8549	0.5245	1.630	0.105806	
DepCaqueta	1.368	0.5177	2.642	0.009368	**
DepCauca	1.512	0.4667	3.240	0.001555	**
DepCesar	1.09	0.4724	2.306	0.022849	*
DepChoco	0.855	0.5144	1.662	0.099164	.
DepCordoba	0.8360	0.4698	1.780	0.077739	.
DepCundinamarca	0.1404	0.3463	0.405	0.685914	
DepGuajira	1.09	0.4809	2.266	0.025306	*
DepHuila	1.571	0.4746	3.311	0.001238	**
DepMagdalena	0.8084	0.4888	1.654	0.100863	
DepMeta	0.4059	0.4463	0.909	0.365030	
DepNariño	1.503	0.4770	3.152	0.002063	**
DepNorte de santander	1.204	0.5019	2.399	0.018027	*
DepQuindio	1.127	0.5646	1.996	0.048302	*
DepRisaralda	1.082	0.5062	2.137	0.034668	*
DepSantander	0.2847	0.3304	0.862	0.390521	
DepSucre	0.9365	0.5206	1.799	0.074605	.
DepTolima	1.407	0.4738	2.970	0.003617	**
DepValle del cauca	-0.1591	0.2227	-0.714	0.476534	
PIB	0.000007661	4.767e-06	1.607	0.110744	
IDH	14.21	2.237	6.353	4.20e-09	***
desempleo	-0.02662	0.01006	-2.646	0.009263	**

Table 3: Coeficientes (mean model with logit link)

Signif. codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1

Parámetro	Estimador	Std. Error	t value	Pr(> t)	
ϕ	-3.52186	0.06044	-58.27	1.2e-16	***

Table 4: Coeficientes de phi (precision model with logit link)

El modelo presenta un buen ajuste:

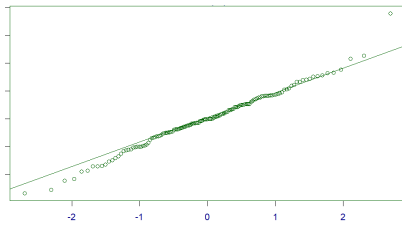


Figure 10: Residuos vs valores ajustados

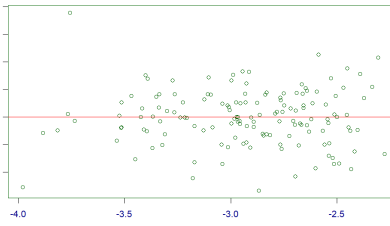


Figure 11: Normal QQ plot

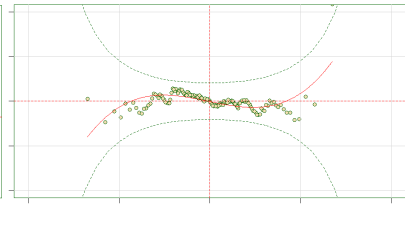


Figure 12: Gráfico de gusano

En las imágenes 10, 11 y 12, se puede apreciar el buen ajuste de los datos, una dispersión aleatoria alrededor de la línea roja sin patrones evidentes (residuos constantes) así como la normalidad de los errores y por último no hay patrones sistemática en la variabilidad de los residuos.

Interpretación de los coeficientes:

- Departamento Bogotá: Debido a que el coeficiente es negativo en comparación con la categoría referencia (Antioquia) la probabilidad de suicidio es menor en Bogotá que en Antioquia.
- Departamento Chocó: Debido a que el coeficiente es positivo, en comparación con la categoría referencia (Antioquia) la probabilidad de suicidio es mayor en Choco que en Antioquia.
- Bajo el modelo beta clásico el PIB no tiene un impacto estadísticamente significativo en la probabilidad de suicidio.
- A medida que el Índice de Desarrollo Humano aumenta, la probabilidad de suicidio aumenta significativamente.
- Un aumento de una unidad en la tasa de desempleo se asocia con una disminución de la probabilidad de suicidio.

4.2.3 Modelo de regresión beta semi-paramétrico

Utilizando la función `gamlss` de Rstudio, se ajustan los modelos de regresión beta que por defecto utiliza la función de enlace "logit".

Se realizó la comparación de los modelos a través de los criterios de información Akaike Information Criterion (AIC) junto con Bayesian Information Criterion (BIC) combinando diferentes variables y/o suavización por splines penalizados.

Modelo	Variables del modelo	AIC	SBC
Mod. uno	Intercepto, Departamento, Año, pb(PIB, IDH, Health, Desempleo, Gini	-998.06	-882.00
Mod. dos	Intercepto, Departamento, pb(PIB, IDH, Health, Desempleo, Gini)	-1001.96	-902.83
Mod. tres	Intercepto, Departamento, pb(PIB, Health, Desempleo, Gini)	-1004.03	-908.06

Table 5: Modelos beta semi-parámétricos ajustados a la tasas de suicidio en Colombia.

Comparando los criterios de información el mejor modelo que se ajuste es el número tres:

$$\text{logit}(\mu_i) = \hat{\beta}_0 + \hat{\beta}_1 \text{Departamentos}_i + \hat{\beta}_2 \text{PIB}_i + \hat{\beta}_3 \text{Health}_i + \hat{\beta}_4 \text{Desempleo}_i + \hat{\beta}_5 \text{Coef.Gini}_i \quad (24)$$

Sobre este modelo se realiza la interpretación del modelo número tres, mostrando las variables significativas (el departamento base fue Antioquia).

Parámetro	Estimador	Std.Error	z.Value	Pr (> t)	
Intercepto	-20.37	2.616e+00	-7.788	3.63e-12	***
DepAtlantico	0.3958	3.650e-01	1.084	0.28056	
DepBogota	-1.270	4.175e-01	-3.042	0.00292	**
DepBolívar	0.3458	3.985e-01	0.868	0.38734	
DepBoyaca	0.8875	4.315e-01	2.057	0.04202	*
DepCaldas	1.012	4.920e-01	2.058	0.04190	*
DepCaqueta	1.135	5.071e-01	2.238	0.02718	*
DepCauca	1.437	4.481e-01	3.207	0.00175	**
DepCesar	1.412	4.428e-01	3.189	0.00185	**
DepChoco	0.8013	4.857e-01	1.650	0.10181	
DepCordoba	0.6619	4.578e-01	1.446	0.15099	
DepCundinamarca	0.1450	3.310e-01	0.438	0.66208	
DepGuajira	2.230	4.648e-01	4.797	4.97e-06	***
DepHuila	1.317	4.655e-01	2.829	0.00552	**
DepMagdalena	1.055	4.602e-01	2.294	0.02367	*
DepMeta	0.2991	4.348e-01	0.688	0.49294	
DepNariño	1.368	4.635e-01	2.952	0.00384	**
DepNorte de santander	1.547	4.653e-01	3.324	0.00120	**
DepQuindio	1.495	5.151e-01	2.901	0.00447	**
DepRisaralda	1.256	4.737e-01	2.652	0.00915	**
DepSantander	0.7105	2.910e-01	2.441	0.01619	*
DepSucre	0.7181	5.104e-01	1.407	0.16224	
DepTolima	1.461	4.516e-01	3.236	0.00159	**
DepValle del cauca	0.1193	1.936e-01	0.616	0.53920	
pb(PIB)	0.00001096	4.356e-06	2.517	0.01324	*
pb(Health)	18.24	2.911e+00	6.267	6.96e-09	***
pb(desempleo)	-0.01598	9.472e-03	-1.687	0.09430	.
pb(gini)	1.103	8.117e-01	1.359	0.17695	

Table 6: Coeficientes (mean model with logit link)
Signif. codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1

Parámetro	Estimador	Std. Error	t value	Pr(> t)	
ϕ	-3.57569	0.06037	-59.23	2e-16	***

Table 7: Coeficientes de phi (precision model with logit link)

El modelo presenta un buen ajuste:

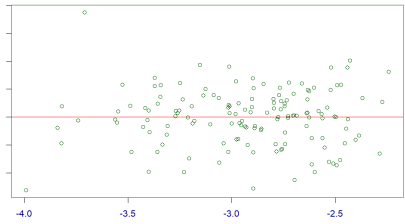


Figure 13: Residuos vs valores ajustados

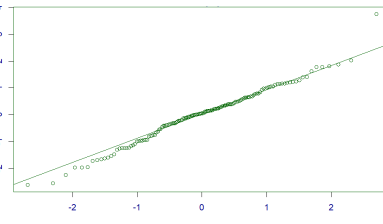


Figure 14: Normal QQ plot

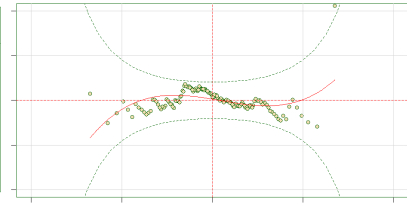


Figure 15: Gráfico de gusano

En las imágenes 13, 14 y 15, se puede apreciar el buen ajuste de los datos, una dispersión aleatoria alrededor de la línea roja sin patrones evidentes (residuos constantes) así como la normalidad de los errores y por último no hay patrones sistemática en la variabilidad de los residuos.

Interpretación de los coeficientes:

- El modelo sugiere que Bogotá tiene probabilidades más bajas de las tasas de suicidio en comparación con Antioquia.
- El modelo sugiere que Tolima tiene probabilidades más altas de las tasas de suicidio en comparación con Antioquia.
- Un aumento de una unidad en el PIB, está asociado con un aumento en las probabilidades de la tasa de suicidio.
- Un aumento de una unidad en la variable de salud, está asociado con un aumento significativo de 18.24 en las probabilidades de la tasa de suicidio.
- Un aumento de una unidad en la tasa de desempleo, está asociado con una disminución en las probabilidades de la tasa de suicidio.
- El coeficiente de gini bajo este modelo no es estadísticamente significativo.

Durante el proceso, se identificaron y analizaron documentos previos que presentaban perspectivas y hallazgos similares. Respecto al PIB, Moyano y Barría (2006) [8] identificaron que en Chile, para los años 1981 a 2003 a pesar de que el país presentó un crecimiento en el PIB, las tasas de suicidio también iban aumentando progresivamente, a pesar de haber un crecimiento económico esto no implica una mejora sino persiste condiciones precarias en la salud mental, precarización laboral, etc.

Otro estudio realizado en México, entre el 2003 y 2007 también evidenció una correlación negativa, además de significancia con otras variables económicas (2019) [9].

4.2.4 Modelo de regresión beta bayesiano

A continuación, se encuentra la tabla número ocho que contiene los modelos beta bayesianos, realizando comparaciones a través del DIC ("Deviance information criterion").

Modelo	Variables del modelo	DIC
Mod. uno	Intercepto, Departamento, Año, PIB, IDH, Health, Desempleo, Gini	-1652.80
Mod. dos	Intercepto, Departamento, PIB, IDH, Health, Desempleo, Gini	-2287.09
Mod. tres	Intercepto, Departamento, PIB, IDH, Health, Desempleo	-2254.89
Mod. cuatro	Intercepto, Departamento, PIB, IDH, Desempleo	-2254.74
Mod. cinco	Intercepto, Departamento, PIB, IDH	-2210.59
Mod. seis	Intercepto, Departamento, PIB, Health, Desempleo, Gini	-2270.36
Mod. siete	Intercepto, Departamento, Health, IDH, Desempleo, Gini	-2086.87
Mod. ocho	Intercepto, Departamento, IDH	-2086.87

Table 8: Modelos beta bayesianos ajustados a la tasas de suicidio en Colombia.

El modelo que presenta el mejor ajuste fue el número dos, escrito de la siguiente manera:

$$\text{logit}(\mu_1) = \hat{\beta}_0 + \hat{\beta}_1 \text{Departamentos} + \hat{\beta}_2 \text{PIB} + \hat{\beta}_3 \text{Health} + \hat{\beta}_4 \text{Desempleo} + \hat{\beta}_5 \text{Coef.Gini} \quad (25)$$

En el modelo se realizaron 100.000 iteraciones descartando 10.000, ejecutando tres cadenas y guardando una muestra de cada cinco. En las tablas 9 y 10 se muestra la estimación de los parámetros estimados utilizando la metodología bayesiana.

<i>Parametrer</i>	<i>Median β</i>	2.5%	97.5%	n.sample	n.effective
Intercepto	-3.372	-7.773	0.736	54000	10
IDH	6.923	0.082	13.719	54000	9
Indice de salud	-0.810	-4.528	2.567	54000	3
Coef. Gini	-0.388	-1.814	1.020	54000	30
Tasa desempleo	-0.019	-0.046	0.010	54000	712
PIB	0.609	0.310	0.899	54000	32
Antioquia	-4.656	-8.778	-1.097	54000	3
Atlántico	-4.369	-8.520	-0.502	54000	4
Bogotá	-6.464	-10.668	-2.998	54000	3
Bolívar	-4.037	-8.173	-0.227	54000	4
Boyacá	-3.424	-7.594	0.465	54000	4
Caldas	-3.149	-7.362	0.911	54000	4
Caquetá	-3.034	-7.232	0.685	54000	3
Cauca	-2.943	-7.139	0.737	54000	3
Cesar	-3.334	-7.560	0.342	54000	3
Chocó	-3.642	-7.882	-0.019	54000	3
Córdoba	-3.647	-7.830	0.007	54000	3
Cundinamarca	-4.126	-8.261	-0.323	54000	3
Guajira	-3.538	-7.732	-0.177	54000	3
Huila	-2.790	-6.984	0.957	54000	3
Magdalena	-3.577	-7.758	0.104	54000	3
Meta	-3.703	-7.907	0.269	54000	4
Nariño	-2.891	-7.062	0.834	54000	3
Norte de Santander	-7.310	-997.23	0.733	54000	4
Quindío	-2.856	-7.110	1.198	54000	4
Risaralda	-3.056	-7.242	0.840	54000	4
Santander	-3.969	-8.109	-0.182	54000	4
Sucre	-3.298	-7.491	0.578	54000	4
Tolima	-2.896	-7.124	0.921	54000	4
Valle del Cauca	-4.480	-8.624	-0.728	54000	3

Table 9: Modelos beta clásicos ajustados a la tasas de suicidio en Colombia.

<i>Parameter</i>	<i>Median β</i>	<i>2.5%</i>	<i>97.5%</i>	<i>n.sample</i>	<i>n.effective</i>
c0	0.507	-13.408	14.074	54000	90
c1	4.019	-13.095	20.972	54000	62
c2	2.730	-13.605	16.528	54000	45
c3	4.140	-11.916	18.530	54000	154
c4	-0.101	-0.422	0.239	54000	273
c5	-0.417	-4.257	3.310	54000	76
c6	3.412	-4.640	11.438	54000	158
c7	-0.339	-4.772	4.155	54000	1864
c8	3.767	-11.207	19.091	54000	89
c9	0.555	-3.834	4.975	54000	260
c10	0.324	-4.102	4.769	54000	244
c11	-0.945	-5.593	3.725	54000	155
c12	-0.808	-5.686	4.113	54000	84
c13	0.492	-3.902	4.946	54000	181
c14	-0.369	-4.756	4.053	54000	154
c15	-1.791	-6.926	3.661	54000	63
c16	-0.388	-4.857	4.088	54000	138
c17	1.319	-3.179	5.794	54000	1296
c18	0.055	-4.853	4.961	54000	182
c19	-1.207	-5.838	3.487	54000	75
c20	-0.288	-4.904	4.323	54000	119
c21	-0.505	-4.913	3.986	54000	194
c22	-0.784	-5.331	3.808	54000	118
c23	-0.052	-4.702	4.692	54000	135
c24	-1.055	-6.129	4.138	54000	180
c25	-0.937	-5.638	3.780	54000	267
c26	0.345	-4.130	4.838	54000	501
c27	-0.903	-5.647	3.885	54000	133
c28	0.196	-4.539	4.989	54000	276
c29	1.451	-4.310	7.178	54000	546

Table 10: Modelos beta clásicos ajustados a la tasas de suicidio en Colombia.

En este modelo se evidencia que las variables significativas fueron el IDH, el PIB y, por último, los departamentos.

Interpretación de los coeficientes:

- El modelo sugiere que el IDH aumenta cuando las tasas de suicidio lo hacen.
- El modelo sugiere que el PIB aumenta cuando las tasas de suicidio lo hacen.
- No se evidenció significancia en la tasa de desempleo, coeficiente de Gini e índice de salud.

4.2.5 Elección del mejor modelo

Con el fin de escoger el modelo se utilizaron tres medidas, desviación absoluta media (MAD), error absoluto medio (MAE) y error porcentual absoluto medio (MAPE).

Modelos	MAD	MSE	MAPE
Modelo beta clásico	0.004921182	0.00004157	9.9382%
Modelo beta semi-paramétrico	0.004597838	0.00003635	9.4785%
Modelo beta bayesiano	0.004621	0.00003926	9.324%

Table 11: Evaluación del mejor modelo

De esta manera se concluye que el mejor modelo entre los descritos anteriormente fue el modelo semi-paramétrico

5 Conclusiones y trabajos futuros

- Existe suficiente evidencia estadística para afirmar que existen factores socioeconómicos que influyen en los suicidios.
- El modelo más robusto se presenta cuando se aplica modelo semiparamétrico siendo significativo el PIB, índice de salud, la tasa de desempleo y los departamentos.
- Para trabajos futuros se sugiere incorporar un modelo beta que tenga en cuenta la información espacial.

6 Agradecimientos

Agradezco a mi tutor Dagoberto Bermudez por su orientación y acompañamiento, a mi familia y amigos por su apoyo constante en el proceso.

7 Referencias

- [1] Cepeda, E. & Gamerman, D. (2005), *Bayesian methodology for modeling parameters in the two parameters exponential family*, ESTADÍSTICA 57(168), 93–105.
- [2] Cepeda, E., Jaimes, Daniel., Marin, M. y Rojas, J. (2016). *Bayesian Beta Regression with the Bayesian-betareg R-Package*. Universidad Nacional de Colombia, Sede Bogotá.
- [3] Cepeda, E. (2001), *Modelagem Bayesiana da variabilidade em modelos lineares generalizado*, PhD thesis, Universidade Federal Do Rio De Janeiro.
- [4] D.J. Spiegelhalter, N.G. Best, B.P. Carlin, and A. Linde (2002). *Bayesian measures of model complexity and fit*, Journal of the Royal Statistical Society series B 64, 583-639.
- [5] Eilers, P. H. C. and Marx, B. D. (1996). *Flexible smoothing with B-splines and penalties (with comments and rejoinder)*. Statist. Sci, 11, 89-121.
- [6] Ferrari, S. and Cribari-Neto, F. (2004). *Beta Regression for modelling rates and proportions*. Universidad de São Paulo, São Paulo - Universidade Federal de Pernambuco, Recife.
- [7] Foster, J., J. Greer y E. Thorbecke (1984), *A Class of Decomposable Poverty Measures*, Econometría, vol. 52, pp. 761-776.
- [8] Moyano Díaz, E., & Barría, R. (2006). *Suicidio y producto interno bruto (PIB) en Chile: hacia un modelo predictivo*. Revista latinoamericana de psicología, 38(2), 343-359.
- [9] Páramo-Castillo, J. D., & Páramo-Castillo, J. D. (2010). *Relación del suicidio con indicadores económicos en México*.
- [10] Rigby, R. A. & Stasinopoulos, D. M. (2005), *Generalized additive models for location, scale and shape*, Journal of the Royal Statistical Society: Series C (Applied Statistics) 54(3), 507–554.
- [11] Stasinopoulos, D. M., & Rigby, R. A. (2008). *Generalized additive models for location scale and shape (GAMLSS) in R*. Journal of Statistical Software, 23, 1-46.
- [12] *Suicidio*. (2021). Organización Mundial de la Salud.
<https://www.who.int/es/news-room/fact-sheets/detail/suicide>